

'Ask not what physics can do for biology, ask what biology can do for physics'

Stan Ulam ;

I.1 *Protein folding problem* : qu'est-ce-que c'est et en quoi est-il important ?

Pourquoi est-il important de résoudre le '*folding problem*', c'est-à-dire de comprendre et prédire la conformation native d'une protéine à partir d'une séquence acides aminés ? Et pourquoi a-t-il inspiré toute une cohorte de biophysiciens ?

Les raisons sont nombreuses. La première est qu'il est important de connaître comment l'état natif peut apparaître en suivant les lois de la chimie et de la physique. Comprendre comment une protéine fonctionne, implique de connaître sa structure tridimensionnelle. Toutefois étudier directement les structures tridimensionnelles des protéines est un processus très long ; c'est pourquoi l'idée d'obtenir la structure tridimensionnelle à partir de la séquence linéaire semble plus simple et plus pratique compte tenu du nombre énorme de séquences disponibles.

Résoudre ce problème de repliement est aussi très important en biotechnologie. En effet, dès lors que l'on sait prédire une structure à partir d'une séquence, on pourra concevoir et prédire de nouvelles protéines *ab initio*. Les applications sont nombreuses : biosenseurs , hormones, agents régulateurs biologiques, applications pharmaceutiques etc. Le problème du *protein design* est appelé « *inverse folding protein* ». Un algorithme de repliement a comme *input* une séquence d'acides aminés et a pour but de prédire comme *output* la structure native, au contraire d'un algorithme de « *inverse folding* » qui utilise comme *input* une structure native et comme *output* un réseau de séquences qui peuvent donner lieu à la structure native.

Le problème du repliement des protéines est aussi relié à celui de l'évolution. En effet, la complexité d'une protéine découle des informations contenues dans la séquence en bases des nucléotides de l'ADN. La compréhension du repliement chevauche alors l'étude de l'évolution moléculaire et de l'origine de la vie.

La complexité des relations séquence –structure a jusqu’à présent limité l’élaboration d’un algorithme qui puisse être valable pour toutes les protéines (Levitt, 1997). Si cette détermination reste encore inaccessible, c’est en partie parce que notre connaissance des processus physico-chimiques qui gouvernent le repliement d’une protéine et l’acquisition d’une structure stable, reste encore insuffisante (Finkelstein , 1997). **L’objectif de cette thèse est de contribuer à mieux comprendre les relations entre séquence et repliement et séquence et stabilité pour une topologie donnée.**

La compréhension complète, de l’acte spontané qui transfère une information linéaire à un objet tridimensionnel demande un grand nombre d’informations, mais surtout la connaissance approfondie de la structure, de la stabilité et de la dynamique de la protéine durant le repliement. Une caractérisation de la structure native et de la structure complètement dénaturée est alors nécessaire, tout comme l’estimation de la différence d’énergie entre ces deux états pour définir la stabilité de la protéine. Suivre la transition et la dynamique du processus de repliement est enfin nécessaire pour déterminer l’énergie de surface de la molécule et aussi estimer quelles sont les étapes et les interactions clés nécessaires conduisant à la structure tridimensionnelle.

I.2 Le repliement des protéines *in vitro* : hypothèses de résolution

In vivo, c’est-à-dire dans le milieu cellulaire riche en molécules et ions, le repliement des protéines implique un important dispositif assurant la catalyse et la régulation de ce processus (Braig.K. 1998 ; Jaenicke 1993), et permettant, en particulier, une adaptation aux contraintes imposées par le milieu cellulaire. Mais les différents facteurs qui sont impliqués dans le contrôle et la localisation du processus ne donne aucune information sur la conformation et sur la façon dont la chaîne polypeptidique se replie. Cette évidence supporte le principe fondamental énoncé par Anfinsen (Anfisen C.B.1973) ‘*The native conformation of proteins is determined by totality of interatomic interactions and hence by the amino acid sequence, in a given environment*’.

L’étude *in vitro* des protéines permet de s’affranchir de la complexité du milieu cellulaire pour ne s’intéresser qu’au repliement de la protéine.

Puisque le squelette de la chaîne polypeptidique de toutes les protéines a une composition chimique identique, comment les chaînes latérales permettent-elles de diriger le repliement ? Le premier problème du repliement est relié au nombre astronomique de conformations possible que peut prendre une chaîne polypeptidique. En effet, si la protéine se replie suivant une recherche au hasard de toutes les conformations possibles, combien de temps prendra son repliement ?

‘*Longer than the life time of the universe*’ répondra Levinthal en 1968 (Levinthal C. 1968), affirmation, appelée Paradoxe de Levinthal, qui a eu un grand retentissement et qui a dominé la discussion sur le repliement des protéines pendant trente ans. En effet, si une protéine est composée de 100 acides aminés et si on suppose que chaque résidu a deux configurations possibles, il y a environ 10^{30} conformations possibles pour la molécule entière. Si seulement 10^{-11} secondes sont nécessaires pour convertir une configuration en une autre, une recherche au hasard prendrait de l’ordre 10^{19} – 10^{11} ans. Il est clair que la nature a trouvé une solution efficace à ce problème, puisque selon les protéines, le processus de repliement s’opère entre 10^{-2} et 10^2 secondes. Levinthal en conclut que les protéines doivent se replier suivant un chemin spécifique et plus rapide. Le corollaire à cette hypothèse (dite ‘cinétique’) est que la forme native d’une protéine ne serait pas la structure thermodynamiquement la plus stable. En effet, rejoindre la structure d’énergie minimale de façon très rapide sont deux conditions qui semblent mutuellement exclusives.

La compréhension du phénomène de repliement a énormément évolué depuis les années 90 et, à l’heure actuelle, deux hypothèses principales se proposent de donner une solution à ce problème :

- l’adoption d’un chemin préférentiel, de repliement, avec la présence d’états intermédiaires. (*‘old view’*)
- conception de l’énergie de surface de la protéine comme un entonnoir : *landscape picture like a funnel*. Ce courant est dénommé : *‘new view’*.

La première solution propose qu’il y ait des chemins préférentiels pour le repliement avec la possibilité pour les protéines de passer à travers des états partiellement structurés, qui peuvent être transitoires ou stables avec une population bien définie. Le mécanisme de repliement est ainsi envisagé comme étant similaire à une réaction linéaire. Si les étapes sont suffisamment spécifiques, seule une petite région de l’espace conformationnel sera explorée

et le paradoxe de Levinthal sera alors évité. Cette conception est confortée par la mise en évidence d'intermédiaires partiellement repliés, qui peuvent se former au cours du repliement et sous certaines conditions de dénaturation. Mais de récentes études expérimentales donnent une description plus complexe du repliement, avec la mise en évidence de chemins de **repliement parallèles**.

Durant les années 1995-1997, une nouvelle vue, combinant approche théorique et expérimentale décrit le folding en terme d'ensembles statistiques (Dobson & Karplus 1999). La mécanique statistique prédit que les états macroscopiques, dénaturés et intermédiaires, sont des distributions ou des ensembles de diverses conformations des chaînes individuelles. Dans ce schéma deux états macroscopiques distincts ne sont pas nécessairement deux ensembles distincts de conformations microscopiques ; ces deux ensembles peuvent se superposer.

Dans cette nouvelle conception, le repliement est défini par Dill comme '*a parallel flow process of an ensemble of chain molecules; folding is seen as more like trickle of water down mountainsides of complex shapes, and less like flow trough a single gully*'. (Dill K. A. and chan H.S. 1997). En d'autres termes, le repliement peut être décrit comme une série de transitions entre des états d'énergie de plus en plus basse, guidant la protéine par un effet d'entonnoir vers l'état conformationnel de plus basse énergie, la conformation native. La chaîne polypeptidique n'explore alors qu'une fraction des états accessibles. Dans cette représentation, les états intermédiaires correspondent à des 'creux' nécessitant, soit que la chaîne polypeptidique 'revienne en arrière', soit qu'elle se replie suivant une trajectoire parallèle.

L'espace énergétique est l'énergie libre de chaque conformation en fonction des degrés de liberté comme par exemple les angles dièdres (ψ et ϕ) etc...et chaque conformation est représentée par un point sur l'énergie de surface multidimensionnelle. La FigI.1 (C) montre que à l'origine du repliement il existe plusieurs conformations accessibles ayant la même énergie libre, et que la surface accessible est très large. Lorsque le repliement progresse, l'énergie du système et le nombre de conformations accessibles diminuent avec la formation des contacts natifs. L'entropie du système diminue aussi. La présence de collines correspond à des conformations de haute énergie (état de transition) et celle des vallées à des configurations qui sont plus favorables que les voisines (état intermédiaire). La surface d'énergie libre peut être déterminé par simulation Monte Carlo (Sali, 1994).

La nouvelle vue remplace alors le concept de **chemin d'événements séquentiels** par celui de **l'entonnoir avec événements parallèles**. La nouvelle perspective voit le repliement comme un processus de diffusion où les mouvements de chaque chaîne sont asynchrones et chacun dirigé par les mouvements Browniens à travers différentes séquences de conformations qui amènent, enfin, à la structure native.

En utilisant la métaphore de l'énergie *landscape*, il est possible de comprendre intuitivement les différences entre l'idée de chemin de repliement et l'entonnoir.

Dans la Fig I.1 (A), est représenté ce que on appelle 'paysage du terrain de golf de Levinthal', qui montre clairement qu'une protéine qui cherche sa conformation native, au hasard, peut prendre un temps infini. La Fig I.1 (B), montre qu'un chemin préférentiel amène directement et plus vite à la conformation native. Dans cette représentation, le 'tunnel' peut avoir des vallées et des collines, qui représentent comme déjà discuté des états intermédiaires ou de transition. Mais cette représentation suggère qu'au début du repliement il existe une conformation précise A. Or, il n'y a aucune raison, qui implique qu'il existe une conformation de départ spécifique.

La Fig I.1 (C), qui représente l'énergie de surface dans la nouvelle conception, montre tout d'abord que l'hypothèse de rejoindre le minimum de l'entonnoir dans le minimum de temps est possible, mais que, surtout, on peut se trouver en un point quelconque du bord de la surface d'énergie. Ceci signifie que le minimum d'énergie peut être rejoint par plusieurs chemins.

La solution serait donc l'entonnoir' et non les '*tunnels*' (ou chemin préférentiels). La distinction entre le chemin préférentiel et l'entonnoir est que le premier est un chemin à une dimension, à travers l'espace des configurations, alors qu'un entonnoir décrit plutôt une diminution progressive de la dimensionnalité de l'espace accessible, en passant de plusieurs degrés de liberté pour la chaîne dénaturée à une absence complète pour l'état natif.

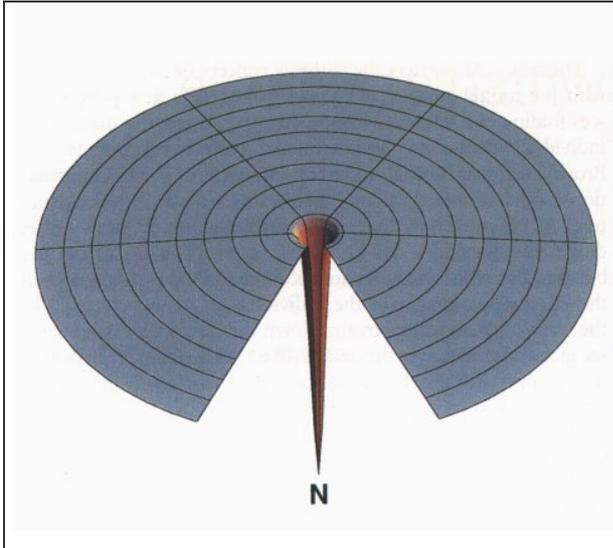


Fig I.1 A) 'Paysage du terrain de golf de Levinthal. *N* est la conformation native (Dill and Chan. 1997).

Fig I.1. B) *Le chemin préférentiel est une solution à la recherche au hasard. A représente la configuration dénaturée* (Dill and Chan. 1997).

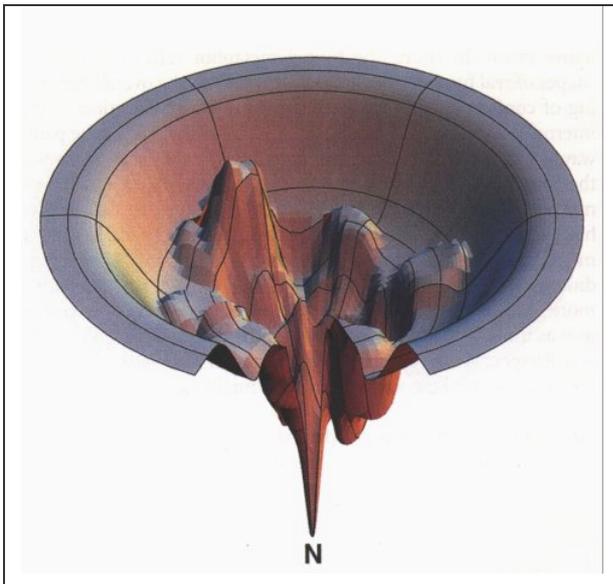
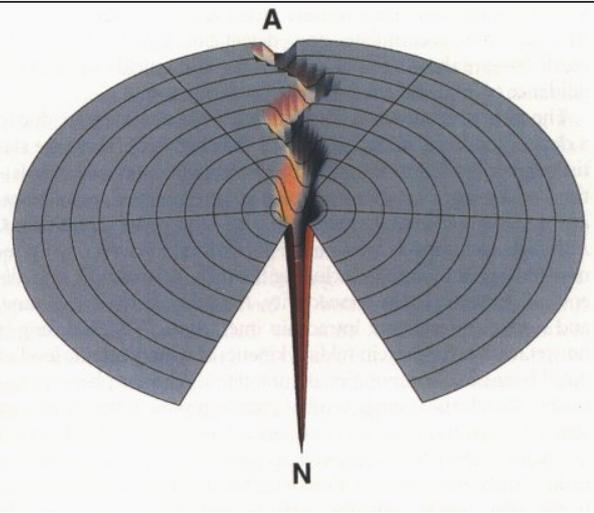


Fig I. 1(C) *Energie de surface avec des pièges cinétiques et barrières d'énergie. Dans cette représentation le repliement peut inclure plusieurs états intermédiaires.* (Dill and Chan. 1997).

I.3 La réaction de repliement.

La *réaction* de repliement est toujours caractérisée par un état initial où la protéine est représentée par une chaîne non structurée et un état final où la protéine atteint la structure native tridimensionnelle. Il est donc très important de bien décrire ces deux formes extrêmes.

Dans le cas plus simple où l'**état natif** est en équilibre thermodynamique avec l'**état non replié** de la protéine (Fersht A.R. 1997 ; Jackson. S.E. 1998.), le processus peut être décrit par un modèle à deux états. Le schéma représentant la réaction est décrit par la Fig I.2

Dans ce modèle, la réaction est une transition du premier ordre dans laquelle les deux états en équilibre sont séparés par une barrière d'énergie correspondant à l'**état de transition**.

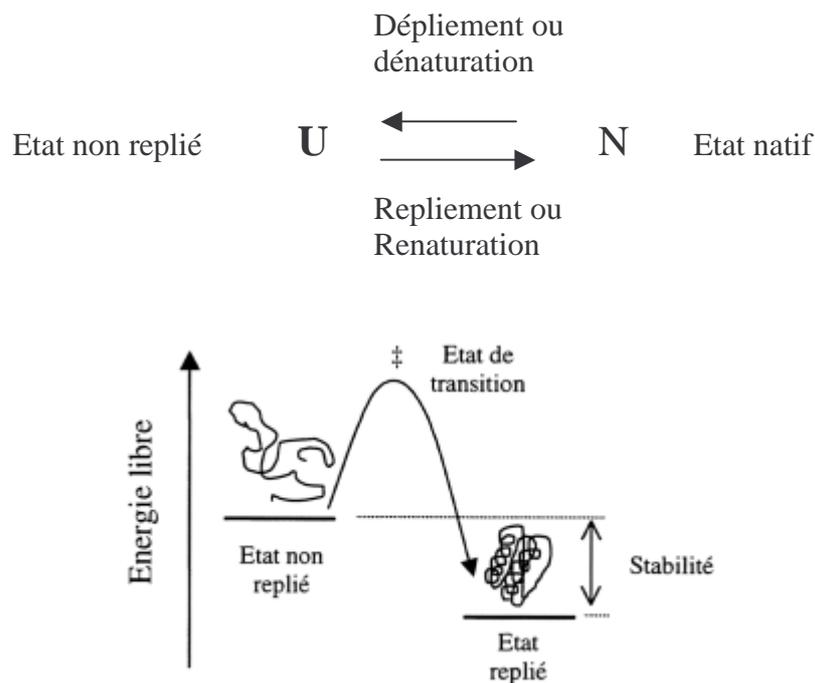


Fig I.2 Schéma représentant la réaction de repliement à deux états dans laquelle l'état non replié est en équilibre thermodynamique avec l'état replié. Ces deux états sont séparés par une barrière d'énergie correspondant à l'état de transition.

Avant de discuter l'étape de transition, nous allons, tout d'abord, décrire l'état natif et l'état complètement dénaturé.

I.4 Caractéristiques de l'état natif

Les techniques de cristallographie et RMN ont permis de déterminer de nombreuses structures de protéines et de décrire les propriétés caractéristiques de l'état natif. La structure générale d'une protéine est déterminée par une hiérarchie d'unités structurales partiellement indépendantes (Fig I. 3)

La **structure primaire** est le niveau d'organisation structural le plus élémentaire et correspond à la séquence en acides aminés. Le second niveau hiérarchique est fourni par les **structures secondaires**, correspondant à un enchaînement répétitif de liaisons peptidique possédant des couples ϕ et ψ particuliers. Ces structures sont essentiellement constituées d'hélices α et des feuillets β . Elles sont connectées entre elles par des boucles situées essentiellement à la surface de la protéine où les contraintes stériques sont minimales. La **structure tertiaire** correspond à l'organisation tridimensionnelle, biologiquement active, ou native de la protéine. D'une façon générale la forme globulaire des protéines est optimisée de façon à enfouir les résidus hydrophobes et exposer les résidus polaires. Enfin un dernier niveau hiérarchique est constitué par la **structure quaternaire**, qui correspond à l'assemblage non covalent de plusieurs chaînes identiques ou non .

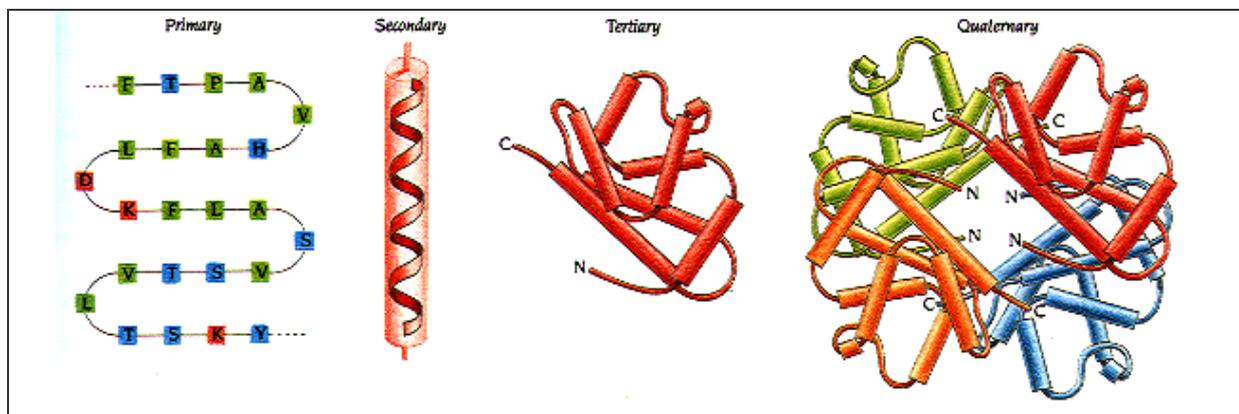


Fig I. 3 Schéma des structures hiérarchiques d'une protéine (Branden 1991)

L'amélioration des connaissances sur les structures des protéines et la continuelle augmentation des structures déposées à la Protein Data Bank ont permis de classer les protéines en dix familles représentatives des topologies les plus fréquentes (Orengo et coll. 1997) Fig I. 4.

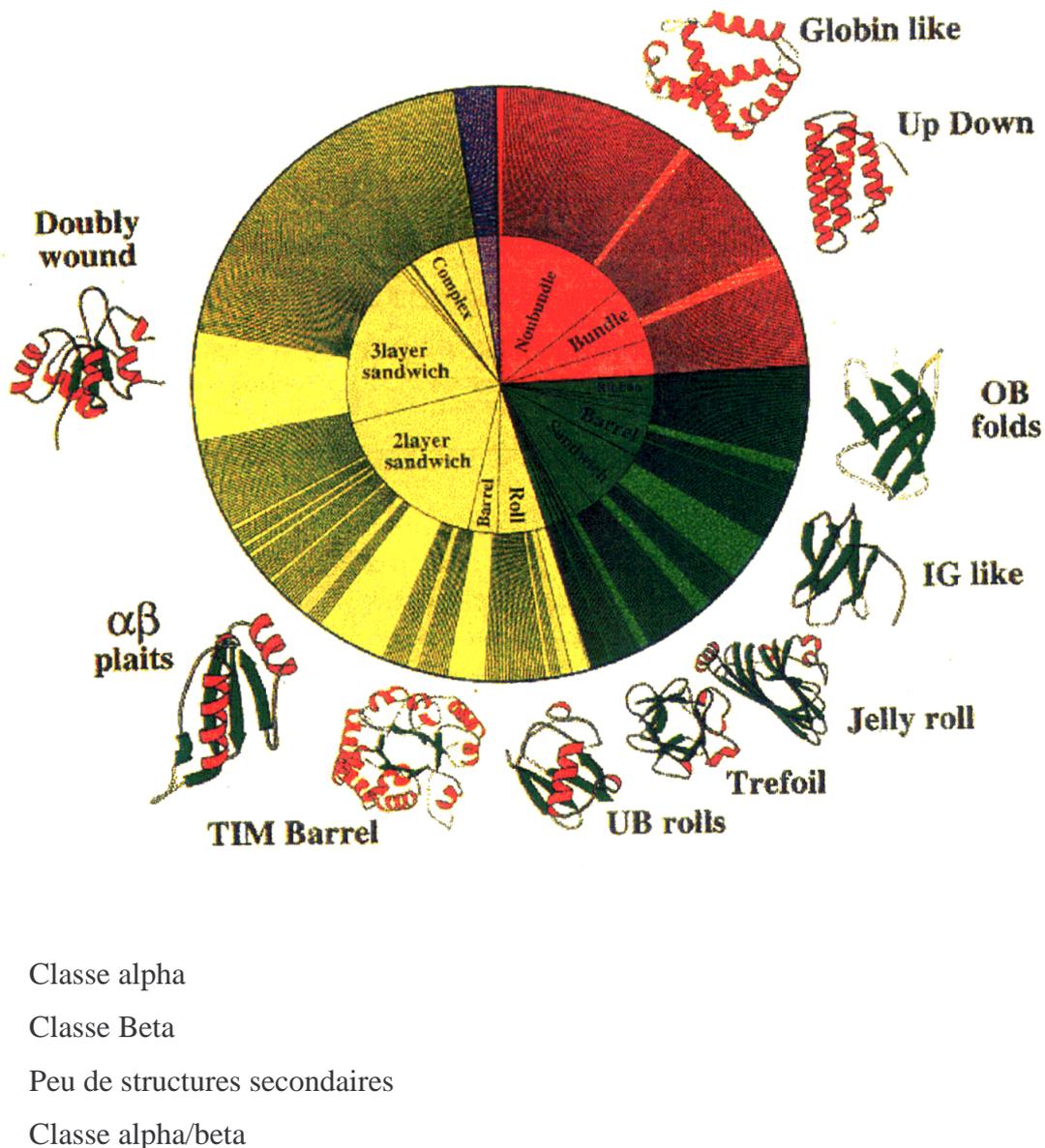


Fig I. 4 La roue CATH représentant le classement des structures de protéines (Orengo et coll, 1997).

On remarque que les séquences des protéines appartenant à une classe topologique donnée ne sont pas forcément identiques et peuvent même être très différentes.

Par contre, Sudarsanam (Sudarsanam 1998) a démontré en analysant 5420 structures de protéines, qu'il existe des segments polypeptidiques ayant une séquence identique qui présentent différentes structures tridimensionnelles. Cela montre que la structure d'une chaîne

est déterminée non seulement par des interactions locales entre les acides aminés proches dans la séquence mais aussi par des interactions non locales.

Pour examiner comment l'environnement influence la formation des éléments de structure secondaire, Minor et Kin (Minor 1996) ont conçu une séquence de 11 acides aminés ('chameleon' séquences), qui se replie comme une hélice alpha ou comme un feuillet beta selon sa position dans la séquence du IgG-binding domaine de la protéine GB1. Plus récemment, Davies *et al* (1998) ont démontré expérimentalement que le peptide de fusion du virus de la leucémie féline (feline leukemia virus fusion peptide) qui est constitué de 28 acides aminés peut changer sa conformation entre un *random coil*, une hélice α ou feuillet β selon l'environnement où il se trouve. Damaschun et coll (1999) démontrent que la protéine phosphoglycerate kinase (PGK), peut atteindre deux autres conformations différentes native, selon l'environnement où elle est placée pendant le repliement. Ces conformations sont moins compactes et contiennent plus de structures secondaires que l'état natif. Ces observations soulignent l'intérêt de l'énoncé de Anfinsen (1973), mais surtout qu'il est très difficile de discerner les résidus ou les groupes d'acides aminés déterminants pour l'acquisition d'une topologie donnée.

L'état natif d'une protéine correspond à un état **thermodynamique stable**, c'est-à-dire à l'état d'énergie libre minimum. Pour le démontrer il est suffisant de démontrer que l'état natif est une fonction d'état qui ne dépend pas du processus de repliement, et donc qu'il constitue l'état d'énergie minimale accessible pendant le processus de repliement. Cette idée a été confirmée par la découverte de la réversibilité du processus de repliement. Les premières expériences sur la réversibilité furent celles de Anson et Mirsky (1931) and Anson (1945) qui démontraient la réversibilité du repliement de l'hémoglobine en mettant en évidence la similarité des propriétés de l'état natif et de l'état 'renaturé'. Depuis il a été montré que la dénaturation de plusieurs protéines, qu'elles comprennent un ou plusieurs domaines est thermodynamiquement réversible. Cela signifie que la structure finale doit correspondre au minimum d'énergie, celui le plus bas de tous les autres états accessibles.

La conformation d'une protéine globulaire résulte d'un équilibre entre les forces dominantes inter-atomiques (interactions hydrophobes, liaisons hydrogènes, interactions de van der Waals, interactions ioniques) et l'entropie conformationnelle de la chaîne

polypeptidique qui s'oppose au repliement (Dill, K.A.1990). Les interactions hydrophobes¹ qui ont été pendant longtemps considérées comme le 'moteur du repliement' ne sont pas la seule contribution dominante. D'abord Tanford (1962), puis Brands (1964) démontrent que si la protéine est conduite au repliement simplement par l'hydrophobie, la stabilité de la protéine serait d'un ordre de grandeur plus grand de celui mesuré. La valeur de l'énergie libre serait de l'ordre de 100- 200 Kcal /mol a 25°C, alors qu'elle est généralement de l'ordre de 5- 10 Kcal/mol (Privalov 1988). Ensuite la tendance au dépliement en fonction de la température n'est pas en accord avec le folding plus '*serré*' qui résulterait théoriquement du comportement des résidus non polaires qui s'associent plus fortement lorsque la température augmente.

Le nombre de conformations compactes accessibles pour une chaîne augmente exponentiellement avec la longueur de la chaîne (Chan & Dill 1989; Dill 1985). Tous les types d'interactions jouent un rôle, dans ces configurations accessibles, et peuvent donner lieu à des conformations ayant des énergies libres différentes. Mais l'existence d'une conformation native unique dérive du fait que les conformations capables de maximiser le nombre de contacts non polaires sont très limitées (Dill, 1990 ; Chan H.S. and Dill K.A. 1996)

Les différents types d'interactions qui définissent la topologie native des protéines peuvent se regrouper en deux grandes classes : **interactions locales et interactions non locales** (Fig I. 5). Une caractéristique très importante du repliement est que des résidus éloignés dans la structure primaire peuvent se retrouver proches dans la structure tridimensionnelle. Les interactions locales, hydrophobes ou polaires, sont celles qui interviennent entre résidus proches dans la séquence. Elles jouent un rôle dans la prédisposition des différents segments de la chaîne polypeptidique à former les structures secondaires. Les interactions non locales sont celles qui interviennent entre résidus éloignés dans la séquence. Elles peuvent être d'origine hydrophobe ou d'origine polaire et favoriser, par exemple, la formation du cœur hydrophobe.

¹ Quand les résidus apolaires sont exposés au solvant, les interactions solvant-solvant sont plus fortes que les interactions solvant-groupes polaires, de sorte que l'eau forme des cages autour de ces résidus apolaires. Quand la chaîne polypeptidique se replie, les résidus apolaires vont se trouver très proches les uns des autres, et vont casser la structure du solvant. La perte de ces interactions, détermine une augmentation d'enthalpie mais aussi une augmentation concomitante du désordre du solvant et donc de l'entropie. Cette variation d'entropie compense largement la variation d'enthalpie de façon à de rendre spontané le processus de repliement de la chaîne hydrophobe à l'intérieur de la protéine.

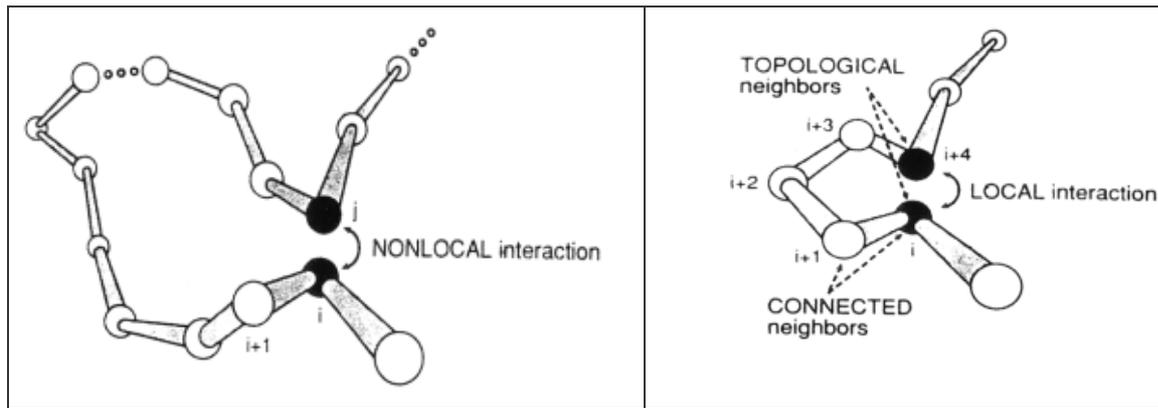


Fig I. 5 Schéma représentant les interactions locales et non locales.(Dill, 1990).

I.5 Caractéristiques des états dénaturés

Le processus du repliement est connu depuis le début du siècle. En 1902, Emil Fischer et Franz Hofmeister concluent indépendamment que les protéines sont des chaînes où les acides aminés sont liés covalamment. Chick et Martin en 1911 ont été les premiers à découvrir le processus de dénaturation et à le distinguer d'un processus d'agrégation. Mais, si en 1925 on associait le processus de dénaturation à l'hydrolyse de la liaison peptidique, ce n'est seulement qu'en 1931 que fut avancée l'hypothèse que la dénaturation des protéines constitue en fait un processus de dépliement (Wu, H. 1931). En 1929 il écrit que '*Denaturation is the breaking up of these labile linkages. Instead of being compact, the protein now become a diffusive structure. The surface is altered and the interior of the molecule is exposed*'. En 1931, il écrit encore : '*Denaturation is disorganization of the natural protein molecules, the change from the regular arrangement of a rigid structure to the irregular, diffuse arrangement of the flexible open chain*' (Wu , 1931).

A l'inverse de l'état natif, les états dénaturés sont beaucoup plus difficiles à définir et à caractériser en raison de l'absence d'une structure précise. Cependant, *les études structurales des ces états désordonnés sont très importantes en raison du rôle des conformations préférentielles initiant le repliement. D'autre part, on ne possède que des informations très fragmentaires sur les événements qui précèdent la formation des structures secondaires et tertiaires.* Si les études de cristallographie sont interdites par l'impossibilité de cristalliser un état dénaturé, depuis quelques années la RMN et la diffusion aux petits angles, des neutrons et de rayons X, ont été largement utilisées, pour obtenir des informations structurales à haute et basse résolution. La RMN est intéressante pour caractériser les structures locales, au niveau

atomique, dans différentes conditions de dénaturation. La diffusion aux petits angles, comme nous en discuterons plus en avant, permet de caractériser la compacité de la protéine et les conformations de la chaîne dépliée.

Les premières études sur les états dénaturés ont été faites par Tanford (Tanford, 1968), qui donne une description détaillée sur la caractérisation de ces états et qui montre surtout comment l'obtenir et comment l'interpréter expérimentalement. Il montre qu'en 6 M et 8 M de GdnHCl, à température ambiante, la plupart des protéines subit une transition vers un état de '*random coil*' sans structure résiduelle, et il souligne une différence de conformation entre '*linear random coil*' et '*cross-linked random coil*' due à l'absence, ou non, de pont disulfure.

Des études plus récentes et en particulier des études de RMN montrent que dans certaines conditions extrêmes de dénaturation, il est encore possible de trouver des structures locales (Damaschun, 1998 ; Belloque, 1998; Buck,1995, Neri 1992). Belloque et coll. (Belloque, 1998) montrent, par exemple que la β -lactoglobuline chauffée à pH 2 conserve la plupart de ses structures secondaires. En étudiant simultanément différentes régions de la protéine à plusieurs températures, ils montrent que si à 55°C certains brins et boucles se déplient complètement d'autres structures ne se déplient que partiellement. Dans 7M d'urée, un amas hydrophobe impliquant quatre chaînes latérales hydrophobes à l'intérieur d'un segment de six résidus a aussi été détecté dans la structure du répresseur 434 (Neri, 1992).

Garcia et al. (1995) ont révélé, pour des concentrations de chlorure de guanidinium proches de 1M, l'existence d'interactions résiduelles autour des deux tryptophanes de la phosphoglycerate kinase (PGK). En travaillant sur des mutants dans lesquels un seul tryptophane est conservé, ils ont caractérisé plus précisément ces structures résiduelles à l'aide de la fluorescence statistique et résolue en temps. Ils ont ainsi montré que ces structures sont stabilisées par des interactions différentes, hydrophobes ou électrostatiques, selon l'environnement local.

L'ensemble de ces études montre que certains états dénaturés sont relativement proches de l'état *random coil* théorique et que certaines structures locales peuvent être formées par la présence de structures secondaires ou amas hydrophobes (Dill and Shortle 1991). Enfin les données expérimentales indiquent aussi que si certaines protéines montrent des caractéristiques similaires dans certaines conditions de dénaturation, l'état déplié varie de façon importante avec l'agent dénaturant (dénaturation en pH acide, thermique et chimique), pouvant ainsi conduire à des processus de repliement différents.

L'état dénaturé comprend un grand nombre de configurations significativement différentes en équilibre rapide entre elles. Il peut être donc défini comme une distribution de différentes configurations qui dépend de facteurs externes. Sous certaines conditions **l'état complètement déplié**, c'est à dire très ouvert sans structures secondaires, peut être prédominant (Dill 1991). Les **états dénaturés compacts**, sont obtenus dans des conditions de dénaturation plus douces Fig I. 6.

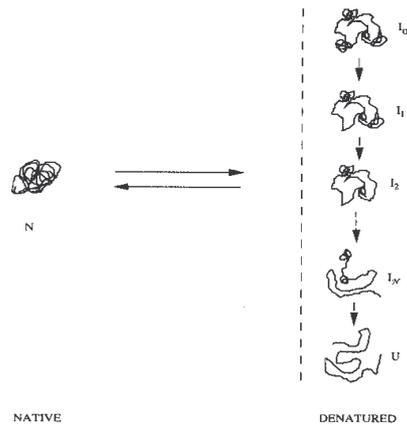


Fig I. 6 Schéma de la transition de dénaturation. A gauche est représenté la protéine native ; à droite les conformations possibles de la protéine dénaturée. I_0 représente un état compacte possible. U correspond à l'état complètement déplié. (Griko, 1994).

L'augmentation du nombre d'études sur les états dénaturés, compacts et dépliés, a stimulé le développement des méthodes d'analyse et interprétation des données expérimentales souvent très complexes. Par exemple, une stratégie pour détecter les structures résiduelles consiste à utiliser la RMN et à interpréter les déplacements chimiques, des protons, par rapport aux déplacements chimiques que l'on attend pour ces protons dans un *random coil*. Fiebig et al (Fiebig, 1996), proposent une stratégie plus complète, pour décrire les conformations du squelette des protéines dénaturées. Ils utilisent la distribution des angles dièdres de la chaîne principale qui est extraite des bases des données. Ce modèle, qui est utilisé pour prédire les paramètres de la RMN, reproduit plusieurs caractéristiques des données expérimentales pour le lysozyme dans 8M d'urée à pH acide. Ceci permet l'identification des structures résiduelles, pour les états non natifs des protéines. Les structures résiduelles observées dans le lysozyme peuvent alors être expliquées par des changements locaux des angles dièdres induits par la variation de l'état d'ionisation des résidus. Une autre

possibilité est que les structures résiduelles peuvent être déterminées par des effets coopératifs entre les résidus proches de la séquence.

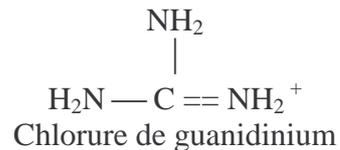
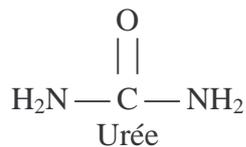
1.5.1 Obtention des états dénaturés.

Les modèles expérimentaux de l'état dénaturé sont obtenus à partir de perturbations de l'état natif, et peuvent être obtenus de façons très différentes : par variation de la température, du pH et par addition d'agents chimiques dénaturants.

La variation de pH en modifiant l'ionisation des chaînes latérales de la protéine, déstructure les paires ioniques et rompt les liaisons hydrogènes. En général les protéines se déplient pour des valeurs de pH inférieures à 5 et supérieures à 10. (Creighton 1993)

L'augmentation de la température d'une solution protéique augmente l'énergie de vibration et de rotation des molécules. Au delà d'une certaine amplitude, ces mouvements changent profondément l'équilibre entre les différentes interactions faibles qui donnent sa stabilité à la conformation globulaire de la protéine

L'urée et le chlorure de guanidinium sont les dénaturants chimiques les plus utilisés et les mieux connus.



Les effets directs de ces dénaturants chimiques, sur les protéines, ne sont pas vraiment très bien connus. Ces dénaturants chimiques ont la potentialité de former plusieurs liaisons hydrogènes avec la chaîne peptidique, cassant ainsi les liaisons hydrogènes des protéines et déstabilisant les structures secondaires. Ils ont tendance à augmenter la solubilité des protéines et à interagir de préférence avec la surface, polaire ou non, de la protéine d'une façon plus favorable que l'eau. Ces dénaturants modifient aussi la structure du solvant parce qu'ils ont la capacité de former des liaisons hydrogènes de géométrie différente de celles qui se produisent dans l'eau..

Le problème le plus important est la difficulté de mesurer les constantes de liaison quand la concentration de la protéine est de 10^{-4} M et celle du dénaturant de 4 M (Shellman

J.A. 1987). Il est clair que le dénaturant interagit directement avec la protéine et que à très haute concentration une modification substantielle de la structure du solvant est certaine (Breslow R & Guo T, 1990).

Différents modèles ont été développés pour expliquer comment le dénaturant déplie une protéine. On peut supposer que l'agent chimique en migrant à l'intérieur de la molécule puisse former des liaisons hydrogène avec les atomes du squelette (Hedwig 1991), ou seulement diminuer les effets hydrophobes (Wetlaufer 1964)

Une étude très intéressante a été effectuée par cristallographie aux rayons X, sur des cristaux de protéines (ribonuclease A, dihydrofolate reductase (DHFR)), hydratés par du guanidinium et de l'urée (Dunbar. J. et al 1997). Le changement le plus important qui a été mis en évidence est la diminution de la valeur du facteur thermique donc de la mobilité des atomes. Il est mis en évidence que par exemple la DHFR solvatée avec de l'urée perd la mobilité des atomes polaires des chaînes latérales ; la ribonuclease A traitée avec le chlorure de guanidinium perd par contre la mobilité soit du squelette soit des chaînes latérales polaires et apolaires. La perte de mobilité est justifiée par la présence de liaisons entre le dénaturant et certains sites préférentiels parce que les molécules de dénaturant peuvent former des liaisons hydrogènes entre différents parties de la chaîne polypeptidique. La même réduction de mobilité a été aussi suggérée dans des protéines dénaturées en solution (ribonuclease A, lysozyme, cytochrome *c*) par des expériences de calorimétrie (Makhatadze G.I et al. 1992). Ces expériences ont mis en évidence le fait que la liaison du dénaturant à la protéine est accompagnée par une diminution significative d'enthalpie et d'entropie. Ces résultats montrent que les multiples liaisons hydrogène réduisent la mobilité de la chaîne polypeptidique.

I.6 L'étape de transition.

L'état de transition correspond à la barrière d'énergie qui sépare l'état non replié de l'état replié. Dans l'état de transition, les interactions qui favorisent le repliement, atteignent un seuil critique qui permet de compenser la perte d'entropie associée à la formation des structures natives. C'est un état très instable, et une fois atteint, la protéine peut se replier de façon très coopérative. Par contre si l'état de transition est atteint avec difficulté le repliement de la protéine est susceptible d'être lent.

I.7 Existence des intermédiaires dans les trajectoires de repliement.

Le processus de repliement de la plupart des protéines possédant plus d'une centaine d'acides aminés, présente des cinétiques multiphasiques, indiquant l'existence d'états intermédiaires au cours de leur repliement (Roder et Colon , 1997). Il se pose alors le problème de déterminer les facteurs qui conduisent à la formation de ces intermédiaires.

En comparant les protéines qui se replient avec ou sans la formation d'intermédiaires, on remarque que ni la complexité topologique, ni les vitesses de repliement ni la complexité de l'état de transition ne permet vraiment de les différencier. Il semblerait que le seul paramètre qui distingue ces deux classes de protéines soit la longueur de la chaîne polypeptidique. On observe que la formation d'intermédiaires est associée aux protéines ayant plus de 100 résidus.

Les états intermédiaires obtenus à l'équilibre sont couramment désignés par le terme général de 'globule fondu' ou état partiellement replié. Etant observées à l'équilibre, les propriétés structurales et dynamique de ces états partiellement repliés ont pu être caractérisées en détail.

Les états intermédiaires cinétiques ne peuvent être observé au cours du repliement que s'ils sont suffisamment stables. Les intermédiaires de repliement sont généralement formés rapidement puis évoluent durant une étape plus lente vers l'état natif. Selon les protéines les intermédiaires apparaissent dans l'échelle de temps de quelques millisecondes à quelques centaines de millisecondes.

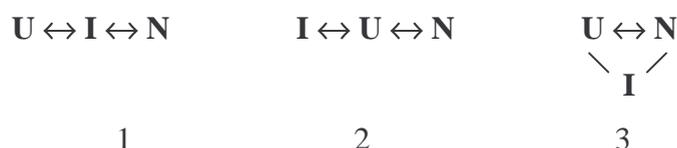
Pour une même protéine, ces intermédiaires peuvent posséder des propriétés très différentes suivant l'état initial de repliement. Les mesures de dichroïsme circulaire résolue en temps, indiquent que les taux d'hélices peuvent varier de 30% du signal natif dans la Snase (Sugawara et coll 1991) à 70% dans la Rnase H (Yamasaki).

Les feuillets β ont été moins bien observés dans ces intermédiaires probablement parce que leur détection est moins aisée par CD. En revanche, de récentes études par spectroscopie infrarouge résolue en temps indiquent que ces intermédiaires peuvent également comporter un taux non négligeable de structures en feuillets (Panick 1998).

Dans l'ensemble, les informations obtenues sur les intermédiaires formés précocement au cours du repliement rendent compte d'un ensemble d'états dynamiques relativement compacts, comprenant un cœur hydrophobe rudimentaire et des éléments de structure secondaire marginalement stable.

1.7.1 Rôle des états intermédiaires.

La question qui se pose est de savoir quel rôle ont les intermédiaires dans le processus de repliement. Pendant de nombreuses années l'hypothèse était que ces intermédiaires correspondent à des phases successives d'organisation de la chaîne polypeptidique aidant à l'acquisition de la structure native. La preuve expérimentale qu'une protéine est capable de se replier suivant un processus à deux états a profondément bouleversé cette représentation car elle a montré que la formation d'intermédiaires stables sur les trajectoires de repliement n'est pas une condition pour l'acquisition de la topologie native (Fersht 1997, Jackson et Fersht 1991). Dans les modèles actuels, les scénarios proposés pour rendre compte du rôle de ces intermédiaires s'articulent autour de trois hypothèses, schématisées par des équilibres chimiques entre l'état non replié U, un intermédiaire I et l'état natif N :



Dans le premier cas l'intermédiaire est une étape obligée du repliement, nécessitant que toutes les molécules passent par cette étape pour se replier. Dans le deuxième cas, l'intermédiaire est un piège cinétique et doit se déstructurer pour conduire au repliement productif de la protéine. Le cas (3), propose un concept supplémentaire qui est celui de l'existence de trajectoires multiples dans le repliement. Dans ce modèle, la chaîne polypeptidique peut emprunter des trajectoires très différentes pour se replier, certaines d'entre elles comportant la formation d'états intermédiaires, d'autre pas.

La présence d'intermédiaires étant détectée pour de nombreuses protéines, la question est de savoir s'il favorisent ou défavorisent le repliement. Plusieurs auteurs ont proposé que l'accumulation d'un intermédiaire ralentit la cinétique de repliement par rapport à celle qu'on pourrait attendre dans un processus de repliement à deux états. Mais certaines expériences

montrent que si l'on mute les résidus participant à la stabilisation d'un intermédiaire, il est possible de passer d'un processus à trois états à un processus à deux états (Baldwin 1996), et que le repliement est plus rapide pour la protéine sauvage, lorsque que l'intermédiaire se forme, que pour la protéine mutée, qui subit une transition à deux états. Rodes et Colon (1997) suggèrent que les intermédiaires formés aux étapes précoces du repliement et impliquant des interactions natives, accélèrent le processus de repliement, tandis que les intermédiaires formés dans les étapes tardives et impliquant des conformations non natives, ralentissent le repliement.

Cet ensemble de données souligne que le rôle des intermédiaires varie selon les protéines considérées. Les intermédiaires sont susceptibles de se former lorsqu'une région plus ou moins importante de la protéine se replie de façon coopérative sans pour autant entraîner le repliement de la protéine entière. Il apparaît donc clairement que les protéines ne se replient pas selon une trajectoire unique, mais que les trajectoires sont multiples, chacune possédant une cinétique propre et ses états de transition. C'est pour rendre compte de cette complexité que plusieurs auteurs, se basant à la fois sur des résultats de simulation théorique du repliement et sur les nombreuses expériences de repliement *in vitro* effectuées ces dernières années, ont proposé le schéma de repliement, utilisant la représentation d'un paysage d'énergie décrit précédemment.

I.8 Les étapes initiales du repliement

Dans la littérature concernant le repliement des protéines il est souligné l'importance de deux processus fondamentaux au cours du repliement, l'un global associé à l'acquisition d'un état compact et l'autre plus local associé à la formation de noyaux de structures secondaires. Ces deux processus jouent un rôle significatif parce qu'ils permettent non seulement à la chaîne polypeptidique de limiter l'exploration de l'espace conformationnel mais également de rejoindre la conformation native.

En 1982 Kim et Baldwin proposent le modèle de la charpente (ou *framework model* qui postule que les structures secondaires de la protéine se forment avant la mise en place de la structure tertiaire. Les structures secondaires forment ainsi une 'charpente' Ce modèle

laisse supposer qu'il n'existe qu'un seul chemin menant de la chaîne polypeptidique naissante à la protéine native (Fig I.7)

Ptitsyn and Rashin (1973) proposent le '*modular assembly model*' dans lequel le repliement commence avec le repliement de différentes régions de la molécule (Fig I. 7)

Ces deux modèles peuvent être combinés entre eux, une protéine peut consister de domaines quasi-independants ou de sous-domaines qui peuvent se replier indépendamment, chacun en accord au *framework model*.

Le modèle du **collapse hydrophobe**, proposé par Dill (Dill, 1990 ; Dill, 1995), suggère que la chaîne polypeptidique subit un effondrement, dû aux acides aminés hydrophobes qui se regroupent en formant un cœur, suivi de la formation des structures secondaires et tertiaire (Fig I. 7).

A partir de la comparaison entre les données expérimentales et les résultats de simulation, il a été possible de postuler d'autres modèles, qui se basent sur la formation de noyaux de structure locale qui vont diffuser et interagir sous l'effet de l'effondrement général de la chaîne polypeptidique.

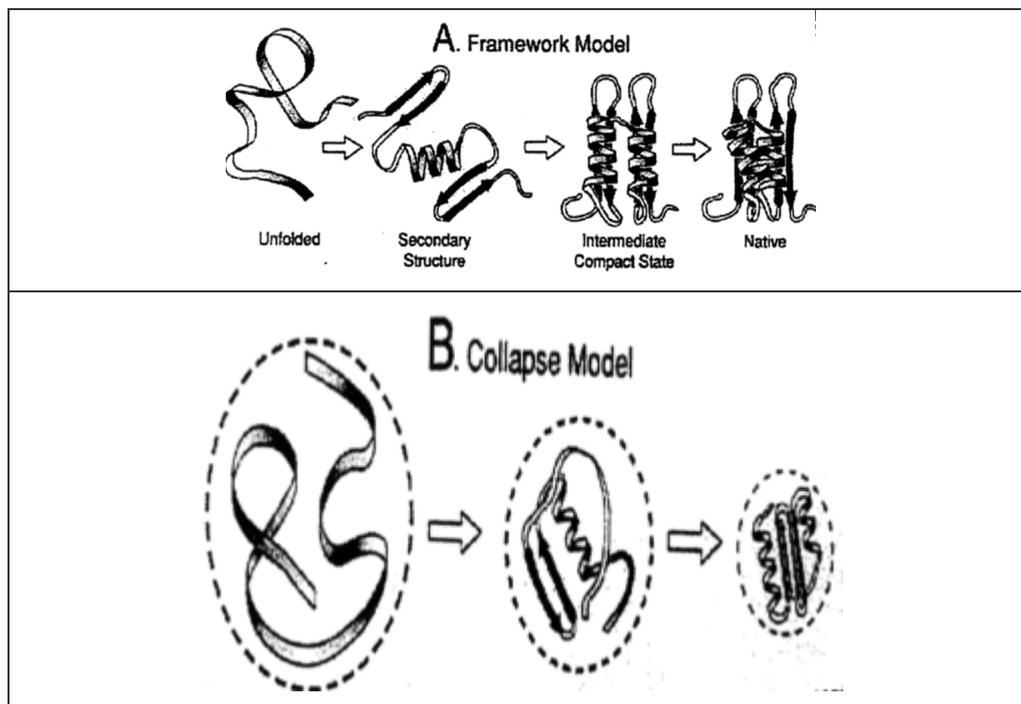


Fig I. 7 Schéma illustrant les modèles de repliement de type charpente (A) et de type collapse hydrophobe (B). (Dill et coll 1995).

C'est le modèle de **diffusion collision** proposé par Karplus et Weaver en 1994, qui postule la formation de microdomaines de repliement qui peuvent entrer en collision par un

processus de diffusion pour former des domaines de tailles supérieures, jusqu'à la structure native. Le modèle de **nucléation condensation** (Itzhaki L.S.,1995) fait aussi partie de ce courant (Fig I.8).

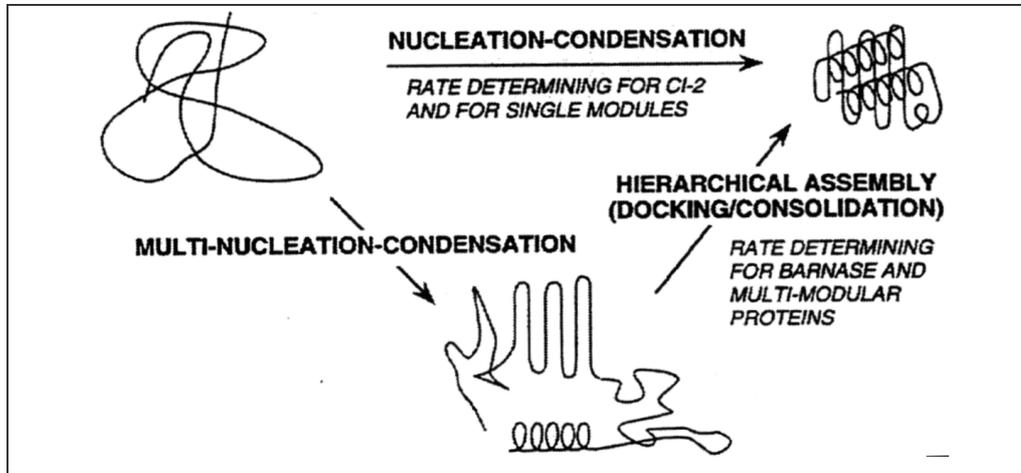


Fig I.8 Schéma illustrant le modèle de nucléation condensation proposé par Fersht 1997.

Le modèle du '*jigsaw puzzle*' proposé par Harrison & Durbin en 1985 [Harrison, S.C. 1985) se différencie des deux premiers car il tient compte de l'hétérogénéité de l'état dénaturé des protéines : puisqu'il y a une grande diversité de conformation des protéines dénaturées, les chaînes polypeptidiques ne devraient pas toutes emprunter le même chemin. Il y aura plusieurs chemins pour arriver à la même solution quasi-unique.

Ces modèles de repliement ont souvent été sujets à polémiques, chacun tentant d'interpréter les résultats expérimentaux en faveur de l'un ou l'autre. Une des raisons de ces polémiques vient du fait que, jusqu'à il y a peu de temps, peu d'informations structurales étaient disponibles sur les intermédiaires observés

Cependant, les études structurales et les développements de la RMN, combinés à la mutagenèse dirigée, permettent maintenant d'identifier de plus en plus souvent les interactions qui stabilisent ces intermédiaires et par suite de proposer un modèle cinétique pour le repliement de certaines protéines. Par exemple, l'étude du cytochrome C, a permis à Roder et ses collaborateurs (1988) de choisir un modèle de repliement séquentiel. Lorsque suffisamment d'informations seront disponibles, il pourra être possible de regrouper différents mécanismes en un seul modèle, exprimant une ligne directrice du repliement, à moins que ces lignes ne soient multiples, à l'instar des structures tridimensionnelles des protéines.

Enfin, des études récentes de mesures cinétiques très résolutive ont permis d'estimer les temps caractéristiques de formation de ces noyaux de nucléation. Ces études montrent que les hélices se forment dans des temps de l'ordre de 0.1 μ s, que la formation de *turns* prends des temps de l'ordre de 0.5 μ s et que la formation d'une boucle qui comprend entre six et dix acides aminés se passe en 1 μ s. Ces données mettent en évidence que les structures locales, codées dans la séquence, se forment très rapidement au cours du repliement et jouent un rôle très important dès les étapes initiales du repliement.

Dans la littérature des dernières années concernant le repliement, il émerge que les petites protéines β semblent suivre de préférence, dans les étapes initiales, soit le modèle de l'effondrement hydrophobe soit celui de la charpente.

Par exemple la formation des structures secondaires intervient dans les premiers pas du repliement de la cobrotoxine, une petite protéine complètement β (7 kDa). Le processus continue avec un *clustering* complète des résidus hydrophobes (Sivaraman, 1999).

Par contre un effondrement hydrophobe intervient dans les premières étapes (< 10 ms) du repliement de la CRABP-I (*Cellular retinoic acid binding protein I*). Cette protéine est réarrangée en 100 à 200 ms en une structure ayant une de type natif (Clark, 1999). Des simulations de dynamique moléculaire sur la protéine betanova (protéine à 3 feuillets beta), ont montré que le repliement de la protéine se déroule en deux étapes d'effondrement. Le premier porte la protéine en un bassin qui contient varies structures qui différent pour la taille et pour les éléments de structures secondaire; à l'état de transition de ce bassin à un bassin native, la protéine adopte une structure et taille native-like et forme une 60% des contacts natives. Avec le deuxième effondrement la protéine acquiert sa complète structure native (Bursulaya, 1999).

Enfin une étude sur la protéine tendamistate, qui a une structure secondaire contenant seulement des feuillets β , des régions avec des boucles et deux ponts disulfure (74 a.a), montre, par des expériences de cinétique en fluorescence, que le repliement se déroule suivant un processus rapide et deux processus plus lents . Le premier, grâce auquel la protéine acquiert 80% de sa conformation native, suit une loi à deux états. Il n'a été détecté aucun effondrement rapide de la chaîne pendant le repliement avant l'acquisition de la structure native tridimensionnelle (Schonbrunner, 1997).

I.9 Techniques biophysiques utilisées pour étudier le repliement des protéines.

Plusieurs techniques développées pour caractériser l'état natif à l'équilibre sont à l'heure actuelle adaptées pour la caractérisation structurale et thermodynamique des états de transition qui apparaissent pendant le repliement. Nous allons résumer dans le tableau suivant certaines techniques qui ont été utilisées au cours des dernières années pour caractériser le repliement des protéines (Plaxco, 1996).

Propriété	Technique	Résolution en temps	Mesure
<i>Core packing</i>	<ul style="list-style-type: none"> Fluorescence Absorbance UV Fluorescence de l' ANS Atténuation de fluorescence 	<ul style="list-style-type: none"> < 1 ms ms ms ms 	<ul style="list-style-type: none"> Orientation et environnement des tryptophanes. Orientation et environnement des tyrosines. Formation et cassure des structures hydrophobes. Enfouissement des tryptophanes
Dimensions moléculaires	<ul style="list-style-type: none"> Anisotropie de fluorescence Diffusion aux petits angles (neutrons et rayons X) Diffusion quasi-élastique de la lumière. 	<ul style="list-style-type: none"> ms <100 ms 1s 	<ul style="list-style-type: none"> Mobilité des tryptophanes. Rayon de giration, longueur de la chaîne. Rayon hydrodynamique.
Structure secondaire et liaisons hydrogène	<ul style="list-style-type: none"> Dichroïsme circulaire lointain UV Pulse labelling NMR Pulse labelling spectrometrie de masse. 	<ul style="list-style-type: none"> ms 5 –1 0 ms 5 –10 ms 	<ul style="list-style-type: none"> Conformation du squelette moyennée sur la séquence. Séquence spécifique de la formation des amides stables et des liaisons hydrogènes. Formation de liaisons hydrogènes.
Structure tertiaire et structure native	<ul style="list-style-type: none"> Activité biologique Repliement interrompu CD proche UV NMR 	<ul style="list-style-type: none"> ms – s 10 ms ms 1 s 	<ul style="list-style-type: none"> Formation de la structure native autour du site actif. Vitesse du dépliement des intermédiaires. Formation des contacts aromatiques et de la formation de liaisons disulfure. Formation de contacts spécifiques de la structure tertiaire.,

I.10 Intérêt de la diffusion des neutrons et des rayons X dans les études sur le repliement des protéines.

Si de très nombreuses structures de protéines à l'état natif ont été obtenues avec précision, par cristallographie ou par résonance magnétique nucléaire (RMN) en solution, les études concernant les états dénaturés n'ont beaucoup progressé que ces dernières années. En effet, la caractérisation structurale de ces états est difficile et les méthodes permettant d'apporter des informations précises sont assez rares (Shortle 1996).

Les récents développements des sources et instrumentations des diffusion des neutrons et des rayons X, ont apporté de nouvelles opportunités dans le domaine de la biologie. En effet, la technique de diffusion aux petits angles permet d'obtenir simultanément des informations sur la taille dans un domaine compris entre 10 et 10^3 Å, la forme géométrique, la compacité et la flexibilité de la molécule ainsi que sur les interactions intermoléculaires.

Dans l'étude du repliement des protéines, la diffusion aux petits angles est une méthode très importante parce que la compacité est le paramètre qui caractérise au mieux le degré de repliement. Depuis quelques années, les travaux faisant appel à cette technique sont en nombre croissant. Ils apportent de précieux renseignements sur les différentes étapes du repliement et sur les interactions stabilisant les divers états dénaturés ou intermédiaires. Grâce à cette technique de diffusion, il a été montré que que les protéines dénaturées pouvaient adopter des conformation très différentes. Les résultats des expériences de diffusion aux petits angles, couplés à ceux de la théorie des polymères (de Gennes 1963 ; des Cloiseaux et Jannink, 1990), ont apporté une contribution importante à la compréhension des propriétés des états complètement dépliés. Cette technique peut être utilisée aussi comme un moyen simple pour contrôler l'intégrité structurale du repliement d'une protéine en fonction des modifications induites par mutagenèse dirigée.

L'application de cette technique à l'étude des protéines complètement dénaturées requiert un certain développement des méthodes d'analyse des résultats. Petrescu (1997) ont proposé une méthode pour calculer le profil de diffusion d'une protéine fortement dénaturée. Dans ce modèle, la protéine dénaturée est représentée comme une chaîne composée d'un

nombre fini de sphères adjacentes. Il a été montré que la similarité de la fonction de diffusion calculée à celle obtenue par la phosphoglycerate kinase dénaturée par 4M de chlorure de guanidinium, augmente asymptotiquement avec le nombre des sphères. Pour décrire l'état complètement dénaturé de cette protéine, 100 sphères (quatre résidus par sphère) sont suffisantes. Le modèle tient compte des interactions locales à courte portée et de celles intervenant entre résidus éloignées dans la séquence. Avec un tel modèle les auteurs ont montré qu'il est aussi possible de décrire des effets de volume exclu (Petrescu A, 1998). Le profil expérimental, de la PGK en 4M de chlorure de guanidinium, peut être représenté par une chaîne flexible de sphères avec des interactions à volume exclu, mais qui peuvent se superposer partiellement.

Konno et al. (1995) ont étudié par les états dénaturés de l'inhibiteur de la subtilisine de *Streptomyces* (SSI) dénaturé par urée, le froid et la chaleur. Ils ont montré que ces trois états sont complètement différents. La protéine dénaturée par 5 M d'urée est un *random coil* dont le déploiement est uniquement limité par les ponts disulfure non réduits. La protéine dénaturée par la chaleur est par contre une chaîne très désordonnée, mais possédant quelques structures secondaires résiduelle, limitant l'extension spatiale de la chaîne. En revanche la protéine dénaturée par le froid, possède une conformation particulière : elle est relativement compacte, une partie de la protéine possédant une structure tertiaire et secondaire différente de celle de la protéine native. L'autre partie est déstructurée et correspond donc principalement, à une région stabilisée à température ambiante par des interactions de nature hydrophobe.

En combinant la diffusion des rayons X aux petits angles et le dichroïsme circulaire Chen et al. (1996) ont mis en évidence un état intermédiaire du lysozyme relativement compact mais possédant une partie *random coil*. Ils ont aussi montré qu'en 8 M d'urée et à pH 2.9, des structures résiduelles persistent encore. Appliquant la méthode analytique SVD (singular value décomposition) ils ont été capables d'obtenir les profils de diffusion de l'état intermédiaire.

La spectroscopie infrarouge (FT-IR), et la diffusion aux petits angles, ont été employées simultanément, par Royer et coll (1998), pour étudier les effets de la pression sur le repliement de la nucléase de staphylocoque. Les résultats montrent que seules des pressions supérieures à 2 kbar ont un effet. A environ 3 kbar le rayon de giration est approximativement

deux fois celui de la protéine native ($R_g=17\text{\AA}$). La variation de la fonction de corrélation de paire montre qu'en même temps, la protéine change progressivement sa forme passant de celle d'un globule, à une forme ellipsoïdale, à celle d'une chaîne étendue. Par ailleurs la FT IR montre que la pression produit des effets dénaturants déjà à partir de 1.5 kbar à (25°C), et que l'état dénaturé obtenu aux pressions supérieures à 3 kbar ne peut pas être décrit comme une simple chaîne aléatoire car des structures β persistent. En fonction de la température la protéine présente une double transition, du rayon de giration, qui atteint sa valeur maximale ($R_g=65\text{\AA}$) à 80°C . A cette température, l'analyse FT-IR suggère que toutes les structures secondaires ont disparu.

Afin d'étudier l'effet de la concentration dans le processus de repliement des protéines Receveur et al (1998) ont utilisé la diffusion des neutrons aux petits angles pour déterminer le deuxième coefficient du viriel de la PGK native et complètement dénaturée. Ils ont montré que la protéine complètement dénaturée par 4M de chlorure de guanidinium présente de fortes interactions à volume exclu.

I.11 La Dynamique des protéines

La modification de la dynamique intramoléculaire, liée aux changements structuraux d'une protéine pendant son repliement ou son dépliement, constitue un problème très complexe à cause de la diversité des mouvements mis en jeu. En effet, la dynamique intramoléculaire d'une protéine a lieu sur un large domaine temporel, qui s'étend de la femtoseconde à la seconde. Les principaux mouvements sont cités dans le tableau 1, avec leurs temps et leurs amplitudes caractéristiques.

Tableau 1 :

Mouvements	Amplitude (nm)	Log ₁₀ du temps caractéristique
Vibrations atomiques	0.001 à 0.01	-14 à -13
Vibrations élastiques de régions globulaires	0.005 à 0.05	-12 à -11
Rotation des chaînes latérales exposées	0.5 à 1	-11 à -10
Rotation des chaînes latérales enfouies	0.5	-4 à 0
Transition allostérique	0.1 à 0.5	-5 à 0
Dénaturation locale	0.5 à 1	-5 à 1

S' il est clair que les changements conformationnels d'une protéine sont très importants pour sa fonction biologique, le rôle des fluctuations thermiques est moins évident (Cusack, Hercules Book). Il a été montré dans certains cas que ces mouvements sont importants pour l'activation de la fonction de la protéine : par exemple, sans fluctuations thermiques, l'oxygène ne pourrait pas entrer dans la poche de la myoglobine . Pour comprendre comment et pourquoi ces mouvements thermiques sont pertinents, une étude de la dynamique dans la gamme de temps où ils interviennent (de la pico- à la nanoseconde) est essentielle.

Les mouvements qui contribuent à la dynamique dans cette gamme de temps sont dus à divers mécanismes, qui peuvent être classés de la manière suivante: mouvements de type vibrationnel, mouvements d'ensemble de domaines, mouvements de type diffusif, en particulier des atomes appartenant aux chaînes latérales des acides aminés. Les vibrations atomiques sont des mouvements rapides, qui correspondent à des énergies supérieures à 50 meV, ou encore à des temps caractéristiques de l'ordre de la picoseconde ou moins (voir l'annexe III pour la correspondance énergie-temps). Les mouvements entre domaines sont attendus dans les protéines pour lesquelles des parties rigides de la structure sont reliées par des parties plus flexibles. Ces mouvements sont suramortis et donc plus lents que les mouvements vibrationnels. Les mouvements des chaînes latérales sont pour la plupart déterminés par leurs interactions avec le solvant, les résidus les plus exposés ayant la plus grande mobilité. Ils interviennent dans l'intervalle d'énergie que nous allons plus particulièrement étudier : 0-50 meV.

L'étude des processus dynamiques au niveau microscopique a été développée en utilisant une grande variété de spectroscopies à haute résolution (spectroscopie infrarouge, spectroscopie Mossbauer, Résonance Magnétique Nucléaire, fluorescence résolue en temps). Il faut rajouter à ces techniques la diffusion quasi-élastique et inélastique des neutrons, qui présente l'intérêt de couvrir des échelles de distances [0.5-10Å] et d'énergies [0-200meV] qui correspondent aux fluctuations thermiques, et de donner des informations sur la géométrie des mouvements. Suivant la gamme d'énergie choisie, il est possible d'étudier les mouvements conduisant à de grands échanges d'énergie, tels que les vibrations atomiques de la molécule, ou les mouvements associés à de faibles échanges d'énergie, correspondant plutôt aux mouvements de diffusion et de rotation des atomes de la molécule. Un autre intérêt de la diffusion quasi-élastique et inélastique des neutrons est de donner accès à des grandeurs que

l'on peut directement calculer à l'aide des simulations de dynamique moléculaire (Deriu 1993).

La première expérience de diffusion quasi-élastique des neutrons a été faite sur des fibres d'ADN dans les années soixante-dix (U.Dahlborg, 1971). Depuis, un certain intérêt s'est manifesté à cause de la capacité des neutrons à caractériser les mouvements de basse fréquence dans les systèmes biologiques. La plupart des études que l'on peut trouver dans la littérature concernent la dynamique de l'état natif. Elles sont focalisées sur l'importance de l'eau d'hydratation (Zanotti, 1999), sur l'étude de la densité d'état vibrationnels (Andreani, 1997 ; Ferrand 1993) et sur la mise en évidence d'une *transition vitreuse* à basse température. Les premières protéines pour lesquelles cette transition a été mise en évidence sont la myoglobine -protéine globulaire- (Doster, 1989 ; Cusak 1990), la bacteriorhodopsine protéine membranaire-- (Ferrand, 1993 ; Fitter 1997, Réat 1997, 1998), la superoxyde dismutase -SOD- (Andreani 1995, Filabozzi 1996). Pour toutes ces protéines, sauf peut être la bacteriorhodopsine, le déplacement carré moyen des atomes présente, en fonction de la température, une transition aux alentours de 180-200K, alors qu'aucune transition structurale n'est observée à cette température. A basse température, les seuls mouvements autorisés sont des vibrations harmoniques de faible amplitude. Au-dessus de la température critique, la dynamique devient de type anharmonique, et elle est caractérisée par des mouvements diffusifs de plus grande amplitude. Cette transition peut être interprétée en terme de sous-états conformationnels, qui correspondent aux différents minima de l'énergie potentielle de la protéine (Fraunfelder, 1991). A basse température, une description harmonique est appropriée, parce que les molécules sont piégées dans un minimum de l'espace conformationnel. A partir d'une certaine température, la protéine possède une énergie suffisante pour franchir les barrières de potentiel. Il apparaît alors des transitions entre sous-états conformationnels, qui permettent à la protéine d'atteindre la conformation dans laquelle elle est fonctionnelle.

I.11.1 Dynamique des états dénaturés des protéines.

La dynamique des états dénaturés des protéines a été étudiée essentiellement à l'aide de la **RMN**, par des mesures de relaxation du ^{15}N . Le nombre d'études est relativement limité car les données sont difficiles à analyser en utilisant les approches standard. Ainsi Buevich et Baum (1999) ont montré que l'approche dite "model free" de Lipari et Szabo, que l'on utilise

pour les protéines natives, n'est plus adaptée pour les protéines dénaturées parce qu'elles sont composées d'un ensemble d'états. A partir de leur travail sur la forme dépliée du pro-peptide de la subtilisine (PPS), ils démontrent qu'il est nécessaire d'introduire une distribution de temps de corrélation, à l'échelle de la nanoseconde, pour analyser correctement les données.

La plupart des études de RMN cherchent à établir une relation entre les fluctuations dynamiques et l'existence au sein de la protéine dénaturée (ou partiellement dénaturée) d'éléments de structure (clusters hydrophobes, éléments de structure secondaire, ponts disulfure). Le premier travail à citer est celui de Alexandru et Shortle (1994), qui ont étudié la dynamique du squelette peptidique d'un fragment en partie déplié de la nucléase de staphylocoque comportant 131 résidus (au lieu de 149). Les auteurs observent que les mouvements sont plus contraints dans les régions hydrophobes, et suggèrent qu'il pourrait exister une relation entre la flexibilité du squelette et la propension à la formation locale d'effondrements hydrophobes, initiateurs du repliement.

Citons également le travail de Wright (1999) sur la structure et la dynamique de deux états partiellement repliés de l'apomyoglobine, et de sa forme native. Il montre qu'une accumulation progressive d'éléments de structure secondaire est suivie par une diminution des fluctuations dynamiques du squelette.

Le lysozyme, dénaturé en 8M d'urée à pH acide, a été étudié en RMN dans sa forme oxydée et dans sa forme réduite. Les résultats ont été comparés aux prévisions théoriques obtenues pour un modèle de type "random coil", basé sur des distributions des angles phi et psi extraites de la PDB. Il y a un bon accord entre les paramètres RMN expérimentaux et calculés, ce qui suggère que les conformations locales des états dénaturés (réduits et oxydés) sont essentiellement déterminées par des interactions à courte portée le long de la chaîne polypeptidique. De plus, l'analyse des données de relaxation met en évidence la formation de "clusters" hydrophobes au voisinage des tryptophanes. Elle montre également que les mouvements sont plus contraints dans les conformations plus compactes qui se forment quand la chaîne est liée par des ponts disulfure.

Enfin, Farrow et al.(1997) ont caractérisé la dynamique du squelette peptidique du domaine N-terminal SH3 de la protéine drk, dans l'état natif et dans un état dénaturé chimiquement, à deux températures différentes. Pour la protéine native, la dynamique du squelette montre peu de variations sur toute la molécule. Par contre l'état dénaturé présente une dynamique plus hétérogène et des amplitudes de déplacement plus grandes. En particulier

les résidus au cœur de la séquence de la protéine sont moins mobiles que ceux des extrémités. Ceci suggère que la molécule ne se comporte pas comme un polymère étendu, et qu'il existe des mouvements concertés dans la partie centrale, compatibles avec une conformation compacte de cette région. Une augmentation de température n'influence que les mouvements de haute fréquence de l'état dénaturé, laissant indifférent l'état natif.

Récemment, des études ont également été menées dans le domaine des *simulations de dynamique moléculaire*. Citons par exemple le travail de Smith L.J. et al. (Smith L.J. et al., 1999) qui ont étudié les modifications structurales et dynamiques de la protéine α -lactalbumine quand elle passe de sa conformation native à pH=6.5, à une conformation partiellement dépliée à pH=2 en l'absence d'ion calcium. Dans la simulation à bas pH, des changements importants sont détectés dans la région C-terminale. De plus, une hélice α (l'héliceD) et l'hélice 3_{10} C-terminale sont fortement perturbées. Simultanément, de larges fluctuations des angles de torsion des chaînes latérales sont observées. Pour les auteurs, ces changements dynamiques et structuraux sont représentatifs du dépliement de la protéine vers la conformation de molten-globule qui la caractérise à pH=2.

La première étude de la dynamique d'une protéine dénaturée par *diffusion quasi-élastique des neutrons* a été initiée par notre équipe, sur la phosphoglycérate kinase -PGK- (Receveur 1997.). Il a été mis en évidence une importante modification de la dynamique de cette protéine quand on la dénature par 1.5 M de chlorure de guanidinium. Dans la protéine native, les mouvements des hydrogènes des chaînes latérales enfouies dans le cœur sont fortement contraints, alors que les hydrogènes des chaînes latérales exposées à la surface peuvent adopter des mouvements diffusifs ayant des temps caractéristiques de l'ordre de la dizaine de picosecondes. Quand la PGK est complètement dénaturée, tous les hydrogènes des chaînes latérales se mettent en mouvement, et seuls les protons du squelette n'adoptent pas de mouvements détectables sur l'échelle de temps explorée au cours de l'expérience [2-30ps]. Cette étude ne donne des renseignements que sur les mouvements internes des protéines native et complètement dénaturée. Nous ne possédons encore aucun renseignement concernant les changements de la dynamique entre ces deux états extrêmes, à cette échelle de temps court.

I.12 Objectifs de la Thèse.

Cette thèse est consacrée à l'**étude des mécanismes de repliement** d'une petite protéine globulaire modèle, la **néocarzinostatine (NCS)**, caractéristique d'un motif structural partagé par de nombreuses protéines : *le repliement des immunoglobines*.

Les similitudes structurales entre la NCS, les domaines des immunoglobulines et d'autres familles de protéines font de la NCS un très bon modèle pour l'étude de la stabilité conformationnelle des structures à **tonneau β** . De plus, cette protéine présente l'avantage d'être très stable, ce qui permet de l'utiliser dans de nombreuses conditions expérimentales.

La stratégie que nous avons employée pour analyser le processus de repliement repose essentiellement sur l'étude de transitions de **dénaturation-renaturation**.

Le principe suivant lequel le mécanisme de repliement d'une protéine puisse être décrypté à partir d'études de transitions de dénaturation-renaturation a été largement vérifié par les études multiples réalisées au cours des 4 dernières décennies. Ceci explique qu'il y a souvent collusion des termes "études du repliement" et "études des transitions", le terme *repliement* ne voulant pas nécessairement dire *renaturation*. À ce titre, la langue anglaise est moins ambiguë puisqu'il existe trois termes distincts (*fold*, *unfold*, *refold*) pour désigner le repliement, la dénaturation et la renaturation.

La première partie de notre étude a eu comme objectif la caractérisation structurale de ces états dénaturés. L'étude approfondie de l'état fortement dénaturé, a été fondamentale afin de caractériser au mieux une des deux références du processus de dépliement.

Ensuite, pour comprendre quel type d'équilibre règle le processus de dépliement, notre étude a consisté à déterminer si le repliement de la NCS suivait une transition à deux états ou permettait la présence d'intermédiaires de dépliement. Enfin, pour apporter des éléments d'information sur la triple relation structure-dynamique –fonction, une comparaison de la dynamique de la NCS native et de la NCS au cours de la dénaturation thermique a été réalisée et nous a apporté des premières informations sur la dynamique d'une protéine partiellement dépliée.

I.13 La néocarzinostatine

La néocarzinostatine (NCS) appartient à une famille de protéines d'origine bactérienne, contenant un chromophore anti-tumoral de type énedyne. Appartiennent aussi à cette famille la C-107, kerarcidine, la macromomycine, l'actinoxatine, et la maduropeptine. A l'exception de la maduropeptine, ces chromoprotéines ont des séquences assez proches : environ 50% des positions des résidus sont conservées dans ces séquences. Cela se reflète dans la similitude de leurs structures (Van Roey, 1982 ; Adjadj 1992)

La NCS fut isolée, pour la première fois, en 1965 du filtrat de *Streptomyces carzinostaticus* (Ishida, 1965) mais la structure du chromophore ne fut élucidée que 20 ans plus tard, en 1985.

L'activité biologique de ces protéines découle directement de la présence du chromophore qui se fixe à l'ADN avec une très haute affinité, provoquant des modifications chimiques irréversibles (Goldberg 1995). La protéine transporte et protège le chromophore de l'hydrolyse (Ishida, 1965).

Le chromophore de la néocarzinostatine contient un cycle à 9 carbones comprenant 2 liaisons acétyléniques conjuguées à une double liaison. Il y a très peu d'informations sur la façon dont le chromophore se fixe à la protéine ou se détache d'elle.

La néocarzinostatine a été isolée et produite au Japon où elle est utilisée pour le traitement de la leucémie, des cancers digestifs (foie, estomac, pancréas, colon) ainsi que des cancers de la vessie. L'utilisation de la NCS est par contre limitée par la toxicité du chromophore et en particulier pour ses effets sur la moelle osseuse. La NCS est aussi très difficile à éliminer par le corps. Si elle est facilement éliminée par le rein, elle est très vite absorbée par les parois de la vessie et remise en circulation.

Un grand nombre de travaux ont été réalisés pour tenter d'améliorer les propriétés pharmacologiques, en modifiant la protéine ou en greffant des polymères biocompatibles à sa surface. L'ensemble de ces travaux s'accorde à montrer que les modifications de la protéine peuvent se traduire par des effets importants sur l'efficacité et la sélectivité de la NCS. Par exemple, un de ses dérivés le SMANCS (stirene-co-maleic acid neocarzinostatine), qui est

une combinaison de deux chaînes polymériques et de NCS, présente une plus grande efficacité clinique, possède une plus grande durée de vie plasmatique et permet une localisation plus sélective vis-à-vis de certaines tumeurs

Ce traitement est désormais approuvé par les autorités japonaises pour le traitement des tumeurs hépatiques (Maeda 1995).

I.13.1 La structure tridimensionnelle.

La première séquence primaire de la NCS fut proposée par Meienhorfer et al (Meienhorfer 1972). Elle fut révisée et complétée par Gibson et al 1984 . La NCS est formée d'une seule chaîne polypeptidique de 113 acides aminés. Sa structure primaire est dominée par les glycines, les valines, les thréonines et les serines qui constituent 60% de ces acides aminés.

La structure tridimensionnelle a été déterminée, par RMN en solution (Adjadj, 1992) et par rayons X avec une résolution de 1.5 Å (Teplyakov, 1993) Fig I.9. La structure du complexe protéine-chromophore a aussi été obtenue par cristallographie avec une résolution de 1.8 Å (Kim 1993).

La structure tridimensionnelle obtenue par cristallographie est en parfait accord avec celles que donne la RMN pour les solutions (Remerowsky, 1990 ; Gao, 1991).

La NCS est une protéine globulaire, composée essentiellement d'un tonneau formé de sept brins β antiparallèles (structure en clé grecque Fig I.9), qui constitue un ensemble très rigide.

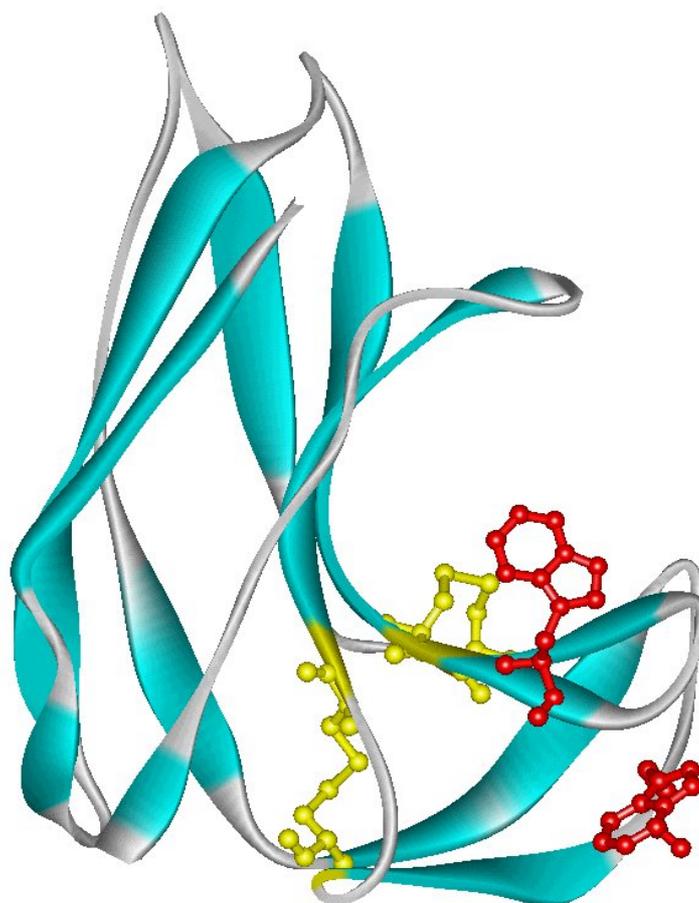


Fig I.9 Structure cristallographique de la néocarzinostatine. Les éléments structuraux en rouge représentent les tryptophanes, en jaune les ponts disulfure et en bleu les feuillets β .

La moitié des résidus, est impliquée dans la structure β . Les éléments de structure secondaire sont donnés dans le tableau I .

La couche interne du tonneau et deux boucles externes forment une cavité hydrophobe qui constitue le site de fixation du chromophore. Cette cavité est définie par les séquences peptidiques suivantes :32-48, 51-52, 73-82, 97-107. Dans la cavité sont présents cinq résidus aromatiques qui interagissent avec la partie aromatique du chromophore.

Une des deux boucles est fermée par les cysteine 37 et cysteine 47 et l'autre par les cysteines 88 et 93. Ces structures présentent une très grande flexibilité en solution, comme le prouve l'analyse par RMN des temps de relaxation du ^{13}C et ^{15}N . Elle suggère que la NCS peut adopter des conformations qui permettent au chromophore de se lier avec des interactions hydrophobes (Adjaji 1992). Parmi les deux ponts disulfure, celui qui relie les Cys 37 et Cys

47 est relativement exposé au solvant. L'autre plus enfoui, est situé à la base du tonneau et est associé aux Cys 88 et Cys 93.

Les tryptophanes de la NCS sont très exposés au solvant.

Tableau 1

Type de structure	Résidus
Trois brins feuillets β	Thr4-Thr8 Val18-Ala24 Ala63-The68
Quatre brins feuillets β	Ala31-Ala38 Ser53-Thr56 Cys93-Ser98 Val108-Ala109
Deux brins feuillets β	Cys37-Asn41 Val44-Cys47
Deux brins feuillets β	Ser72-Phe76 Arg82-Asn87
Coude β II	Ser14-Thr17
Coude β II	Gln27-Thr30
Coude β II	Asn41-Val44
Coude β III	Asp48-Asn 51
Coude β I	Pro49-Phe 52
Coude β I	Asp58-Gly 61
Coude β III	Leu77-Gly 80
Coude β I	Asn87-Thr90
Coude β I	Asp99-Gly702

La structure tridimensionnelle de la NCS est très similaire à celle des autres composants de sa famille dont on connaît la structure : l'actinoxatine (AXN) et la macromomycine (MCR) (Teplyakov, 1993). Le travail de Teplyakov démontre que les structures peuvent être superposées avec une déviation moyenne de la position de 0.129 nm pour les 107 atomes de carbone α communs au couple NCS-AXN et de 0.096 nm pour les 111 communs au couple NCS-MCR. La déviation la plus marquée, se trouve dans les boucles qui caractérisent la protéine et son affinité pour le chromophore associé.

1.13.2 Protéines à “repliement des immunoglobulines”

La famille de protéines à laquelle appartient la NCS se distingue essentiellement par le fait que chaque molécule se lie à un chromophore spécifique. Ces chromophores ont un noyau énédyne commun mais des groupes aromatiques ou glucidiques différents. D'autre part, les chaînes latérales de la protéine qui sont orientées vers l'intérieur de la crevasse et qui sont en contact avec le chromophore ne sont pas conservées dans la famille. Cela suggère qu'une ossature commune soit à l'origine des différentes chromoprotéines et qu'au cours de l'évolution, de nouvelles protéines transportant différents chromophores soient apparues par nouvelles combinaisons de chaînes latérales orientées vers l'intérieur de la crevasse.

Le tonneau β de la NCS présente de fortes similitudes conformationnelles avec le motif structural décrit comme le « **repliement des immunoglobulines** » et qui constitue le corps principal des domaines des anticorps, mais aussi de protéines d'origines et de fonctions très éloignées comme l'hémocyanine ou la superoxide-dismutase (SOD). La SOD possède une fonction catalytique, l'IgG se lie spécifiquement à un antigène par les boucles hypervariables. Par contre la fonction de l'hémocyanine n'est pas encore connue.

Les variabilités observées indiquent que ce tonneau β ne joue pas un rôle fonctionnel direct. Il constitue une architecture structurale stable sur laquelle sont greffées des régions plus flexibles exprimant l'activité biologique de la protéine. Les séquences protéiques donnant lieu à un repliement en tonneau β peuvent ne posséder que 20% d'homologie.

Différentes études comparatives de protéines, qui appartiennent à cette classe structurale contenant un motif en clé grecque, ont été réalisés. Richardson décrit par exemple les similarités entre la SOD et les domaines de IgG [Richardson 1976]. Les variations des séquences de la famille de la SOD ont été analysées par Gestzoff et al. (Gestzoff, 1989) et celles entre les IgG par Bork et al. [Bork,1994].

Adjaji et al [Adjaji, 1992] ont comparé la NCS et le domaine variable Vh du fragment Fab d'une immunoglobuline, par RMN. Les deux protéines, qui ont des fonctionnalités différentes et une homologie de séquence d'environ 20%, présentent une très forte similitude structurale comme le montre la fig I.10. L'homologie du réseau des liaisons hydrogène de la NCS et du domaine Vh suggère que le repliement en tonneau β résulte, principalement, de la formation de liaisons hydrogènes, dans des zones hydrophobes. Dans les

deux structures le tonneau est fermé par une liaison hydrogène entre le C-terminal du 7^{ème} brin et le dernier acide aminé du 1^{er} brin.

Seulement deux différences structurales importantes peuvent être observées : la première est la présence dans le Vh d'un pont disulfure qui lie le 2^{ème} et le 6^{ème} brins et qui relie les deux feuillets du tonneau β passant à travers le tonneau ; la deuxième est une connexion différente des divers brins au sein du tonneau β pour former la structure en clé grecque (Fig I.10).

Une étude comparative des hémocyanines, des immunoglobulines, des superoxide-dismutases et de l'ancinoxanthine, a été réalisée par Hazes et al [Hazes, 1992]. Cette étude suggère que la structure en clé grecque est due à un événement de nucléation universelle, qui ne dépend pas fortement de la séquence, et propose qu'une structure de type ' β zipper', soit importante dans le processus de repliement. Cette étude vient compléter des études précédentes réalisées. Richardson et al.[Richardson, 1976], suggèrent la possibilité que ce motif structural représente une unité de repliement énergiquement favorable et que le grand nombre des résidus présents dans les structures β , peut être le facteur qui contrôle le processus de repliement.

Dans ce model, le β zipper définit le site de nucléation. Il est caractérisé par un grand nombre des résidus hydrophobes, distribués dans deux brins β (B et C qui sont les éléments les plus homologues des différentes structures) (figI.11). Ils se trouvent au centre des feuillets et sont liés par une petite boucle. Cette structure n'est pas maintenue par des liaisons hydrogènes mais par effets hydrophobes. La présence d'une telle structure est conservée dans les quatre familles étudiées et aussi dans d'autres protéines (azurines, plastocyanines etc..) qui ne possèdent aucune similitude avec les immunoglobulines.

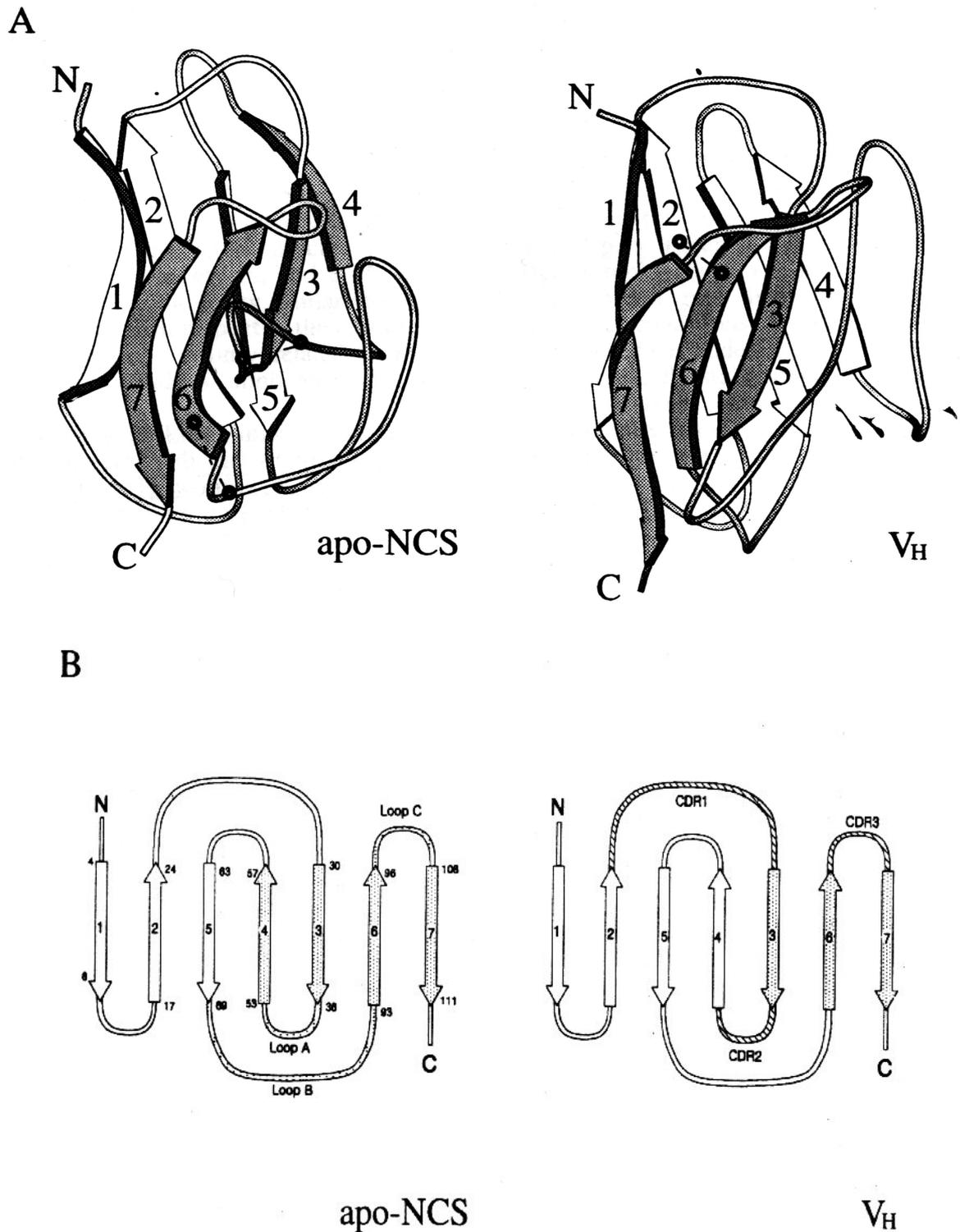


Fig I.10 (A) Comparaison entre les structures du domaine variable des immunoglobulines et de la NCS. Les deux protéines ont les mêmes repliements, La position des cystéines est indiquée par des anneaux. (B) Diagramme topologique de chaque protéine représentant le repliement à clé grecque. Les boucles A, B, C pour la NCS et CDR1 CDR2 et CDR3 pour le domaine variable sont impliquées dans le site actif (Adjaji, 1992).

La comparaison des séquences et des différentes propriétés des acides aminés en positions équivalentes, montre que la présence des résidus hydrophobes dans le coeur est aussi conservée, dans ces quatre familles, comme l'est aussi le volume total des chaînes latérales de ces résidus. Par contre, la distribution des résidus hydrophiles et de la longueur des boucles varie de structure en structure.

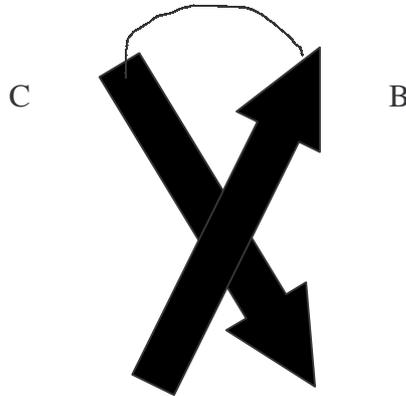


Fig I. 11 Dessin schématique du β zipper. (Hazes ,1992)

Le concept de ‘noyau de repliement’, implique que le repliement commence par la formation d’un ou plusieurs noyaux contenant des acides aminés proches dans la séquence linéaire. Une raison pour cela est que le temps qu’il faut pour créer un ‘noyau de repliement’, augmente avec la distance entre les composants du noyau (Karplus, 1976). La seconde raison est que la perte d’entropie résultant de la fixation de deux segments de peptides est également proportionnelle au nombre de résidus entre les deux segments (Chan, 1988) Le β zipper répondant à ces deux critères pourrait bien participer à un mécanisme de nucléation commun à tous les tonneaux β .

Si on considère alors que cette structure est le site de nucléation du tonneau β , les auteurs proposent un mécanisme de repliement où : 1) toutes les étapes cherchent à optimiser les interactions hydrophobes ; 2) les éléments qui peuvent former des liaisons hydrogènes, ne s’enfuient pas dans la matrice protéique, avant d’avoir formé la liaison H ; 3). les résidus qui sont proches dans la séquence au β -zipper, sont favorisés et adjoints en premier.

De façon plus générale Russel et al. (Russel, 1994), en comparant environ 607 couples de protéines de structure tridimensionnelle similaire suggèrent que chaque type de structure possède différentes interactions stabilisantes qui lui sont propres. Ils proposent que c’est une caractéristique plus générale de la structure des protéines, au même titre que l’énfouissement des résidus hydrophobes ou l’exposition des résidus polaires. Ces hypothèses ont été étayés

par les études de *protein design* où il a été possible d'obtenir des repliements caractéristiques à partir d'un certain nombre d'acides aminés en se basant essentiellement sur leur hydrophilie / hydrophobie (Cunningham 1997 ; Nygren 1997).

I.14 Plan de travail.

L'absence d'études précédentes sur les caractéristiques de repliement de cette protéine nous a obligé à planifier une étude longue et détaillée aussi bien sur l'état natif, l'état dénaturé que sur les transitions obtenues par dénaturation, chimique et thermique.

Afin d'obtenir des informations à différents niveaux structuraux, nous avons réalisé ces études à l'aide de plusieurs techniques. Pour suivre les variations de la structure tertiaire et secondaire de la protéine, pendant le dépliement, nous avons utilisé les techniques classiques de fluorescence et dichroïsme circulaire. A l'aide de la diffusion aux petits angles, neutrons et rayons X, nous avons caractérisé la variation de la taille et de la compacité de la NCS au cours de la dénaturation. La caractérisation thermodynamique, de la transition de dénaturation chimique et thermique, a été réalisée par des mesures de microcalorimétrie différentielle.

Rappelons que tous les signaux spectroscopiques traduisent une moyenne sur les populations présentes pendant le processus de dépliement. De ces considérations découle l'importance d'utiliser un ensemble de techniques, qui fournissent des renseignements complémentaires. Seule la comparaison entre les différentes grandeurs obtenues permet d'avoir une vision plus précise du processus de dépliement et d'obtenir de précieux renseignements parfois cachés par la moyenne des signaux. Par exemple une courbe de transition, de forme sigmoïdale, peut très bien correspondre à une moyenne entre plusieurs transitions ou représenter un phénomène de dépliement continu.

La première étape de notre étude a consisté à caractériser l'état natif et l'état complètement déplié de la NCS. La conformation native en solution a été étudiée en déterminant sa compacité et les interactions intramoléculaires dans différents solvants. Une description détaillée des états dénaturés et des possibles conformations qu'ils peuvent acquérir, a été achevée à l'aide de la diffusion de neutrons et rayons X aux petits angles et de la théorie des polymères. Cette nouvelle approche nous a permis d'ouvrir une porte sur un

domaine difficilement accessible par les techniques classiques. Dans l'étude de la protéine complètement dépliée nous nous sommes intéressés à la réponse de la matrice protéique en conditions extrêmes de hautes températures et hautes concentrations de chlorure de guanidinium. Une différence remarquable de comportement de la chaîne polypeptidique a été mise en évidence et bien analysée au cours du *troisième chapitre*.

Dans le *quatrième chapitre* nous avons décrit, à l'aide des différentes techniques, la transition de dénaturation thermique et chimique de la NCS. La première caractéristique qui émerge est la grande stabilité structurale de la protéine, probablement liée, aussi, à la présence de deux liaisons disulfure. La diffusion aux petits angles nous a permis d'apporter des informations sur la compacité de la molécule au cours de la dénaturation, par la mesure du rayon de giration. Le contenu en structure secondaire et tertiaire a été mesuré sur la protéine aux différents degrés de dénaturation. La comparaison de ces informations, soit pour la dénaturation chimique soit pour la dénaturation thermique, a permis de mettre en évidence l'existence en solution d'états intermédiaires. Une tentative de caractérisation de ces états est présentée au sein du même chapitre.

Enfin, pour compléter la description structurale de la protéine au cours du dépliement thermique, une étude de la dynamique intramoléculaire de la NCS, à l'échelle de la picoseconde, a été réalisée. Le *cinquième chapitre* est dédié à la description de cette étude, qui a été réalisée à l'aide de la diffusion quasi-élastique des neutrons.